

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2003, 23:55:25 ; Search time 6291 Seconds

(Without alignments)
9918.362 Million cell updates/sec

Title: US-09-698-781-2

Sequence: 1 tgatgaacaatactactcat.....gaaaaaaaaaaaaaaaaa 2144

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBank: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rnd: *
36: em_htg_rnd: *
37: em_htg_vit: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2144	100.0	2144	6 AX127587	AX127587 Sequence
2	2134.8	99.6	2138	9 HSPC28	X94323 H. sapiens m
3	2133.2	98.6	2128	6 AX35634	AX35634 Sequence
4	2113.2	98.6	2128	2 HSCRSPPG	X95240 H. sapiens m
5	1485.2	69.3	151752	2 AC010779	AC010779 Homo sapi
6	1480.4	69.0	118524	9 HSDJ44216	AL121950 Human DNA
7	515.2	24.0	1335	9 HUMTPX1A	M25532 Human testis
8	513.6	24.0	1380	9 BC022011	BC022011 Homo sapi
9	513.6	24.0	1406	9 HSCRSPP2I	X95239 H. sapiens m
10	505.8	23.6	1295	4 ECRIISP3	AF001400 Equus cab
11	476.2	22.2	1388	10 CPUS3712	U35712 Cavia porce
12	370.8	17.3	1418	10 MSTRPX1A	M25533 Mouse testis
13	367.8	17.2	1280	10 AB009662	AB009662 Rattus no
14	362.2	16.9	1432	10 AF078552	AF078552 Rattus no
15	362.2	16.9	1498	10 RATAEG	M31173 Rat epididym
16	346.4	16.2	1380	10 MUSAE61A	M92849 Mouse acid
17	346.4	16.2	1403	10 MUSCRISPA	L05559 Mouse cyste
18	341.6	15.9	1445	10 BC011150	BC011150 Mus muscu
19	325.6	15.2	928	10 RNEGR	X04643 Rat mRNA fo
20	325.6	15.2	495	4 ECA6632	AJ006632 Equus cab
21	284.4	13.7	453	4 BTA277708	AJ277708 Bos tauru
22	287.6	13.4	1394	10 MUSAE62A	M92850 Mouse acid
23	287.6	13.4	1418	10 BC022573	BC022573 Mus muscu
24	287.6	13.4	1406	10 MUSCRISPB	L05560 Mouse cyste
25	254.4	11.9	256	6 AX127586	AX127586 Sequence
26	226.4	10.6	1090	5 U13619	U13619 Heloderma h
27	210.8	9.8	334	5 BTA277709	AJ277709 Bos tauru
28	205	9.6	1336	5 AF384218	AF384218 Agkistrod
29	202.4	9.4	1321	5 AF384219	AF384219 Trimeresu
30	198.6	9.3	791	5 AY093955	AY093955 Rhadodophl
31	198.6	9.3	1318	5 PMU59447	U59447 Trimeresuru
32	183	8.5	51371	9 AL353193	AL353193 Human DNA
33	182.8	8.5	169189	2 AC124908	AC124908 Equus cab
34	182.4	8.5	1309	5 AF190861	AF190861 Lepemus h
35	182	8.5	1475	4 ECA315379	AF315379 Equus cab
36	180.8	8.4	1305	5 AF384220	AF384220 Laticauda
37	179.8	8.4	155869	2 AC023420	AC023420 Homo sapi
38	177.4	8.3	1316	5 AF159541	AF159541 Lepemus h
39	177	8.3	1281	9 HUMARP	D38451 Homo sapien
40	177	8.3	1282	9 S80310	S80310 Homo sapien
41	177	8.3	1886	9 HSCRSPP1G	X95237 H. sapiens m
42	175.6	8.2	1174	4 ECA6631	AJ006631 Equus cab
43	173.4	8.1	149907	4 AC091436	AC091436 Felis cat
44	170.2	7.9	166287	9 AC026843	AC026843 Homo sapi
45	169.2	7.9	1343	9 AF123894	AF123894 Macaca mu

ALIGNMENTS

RESULT 1
AX127587
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

AX127587 2144 bp DNA Linear PAT 15-MAY-2001
Sequence 2 from Patent WO0131343.
AX127587
AX127587.1 GI:14134279

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2144)
Hubert,R.S., Raitano,A.B., Afari,D.E., Mitchell,S.C., Faris,M. and
Jakobovits,A.
Diagnosis and therapy of cancer using sgp28-related molecules

JOURNAL Patent: WO 0131343-A 2 03-MAY-2001;
 urogenesys, Inc. (US)
 Location/Dualifiers
 FEATURES 1. 2144
 Source /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 735 a 403 c 382 g 624 t
 ORIGIN

Query Match 100.0%; Score 2144; DB 6; Length 2144;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TGATGAACAATACTATCTCTCTGCTGGAACACCTGCAATGACATTATCCCACTGC 60
DB 1 TGATGAACAATACTATCTCTCTGCTGGAACACCTGCAATGACATTATCCCACTGC 60
QY 61 TGTGTTCTGCTGCTGCTGCTGCTTCCATCTTTCCAGCAATGAGATAGATCCCG 120
DB 61 TGTGTTCTGCTGCTGCTGCTGCTTCCATCTTTCCAGCAATGAGATAGATCCCG 120
QY 121 CTTTACTGCTTTGTTAAACACCAACAAGTCAAGGAGATTGTGAATTAAGCACA 180
DB 121 CTTTACTGCTTTGTTAAACACCAACAAGTCAAGGAGATTGTGAATTAAGCACA 180
QY 181 ATGAAGTGAAGAGACAGATATCTCTCCCTGCGCAAGAACATGCTGAAGATGAATGACACA 240
DB 181 ATGAAGTGAAGAGACAGATATCTCTCCCTGCGCAAGAACATGCTGAAGATGAATGACACA 240
QY 241 AAGAGGCTGACAGCAATGCCCCAAAGTGGCGCAAAACAGTGCATATACAGACAGTAAAC 300
DB 241 AAGAGGCTGACAGCAATGCCCCAAAGTGGCGCAAAACAGTGCATATACAGACAGTAAAC 300
QY 301 CAAAGGATGCAATGACAAAGTCTAAATGTGTGAGATCTCTACATGTCACAGTGCACCA 360
DB 301 CAAAGGATGCAATGACAAAGTCTAAATGTGTGAGATCTCTACATGTCACAGTGCACCA 360
QY 361 GCTCATGCTCACAAGCAATCCAAAGCTGTTGATGATGATACATATTTTGACTTGGTG 420
DB 361 GCTCATGCTCACAAGCAATCCAAAGCTGTTGATGATGATACATATTTTGACTTGGTG 420
QY 421 TAGGGCCAAAGACTCCCAAGCAGTGTGACATTAACAGGTTTGTGACTCTT 480
DB 421 TAGGGCCAAAGACTCCCAAGCAGTGTGACATTAACAGGTTTGTGACTCTT 480
QY 481 CATACCTGTTGAGTGGAAATGCTCTACTGTCCCAATCAAAAAGTCTTAAATACTCT 540
DB 481 CATACCTGTTGAGTGGAAATGCTCTACTGTCCCAATCAAAAAGTCTTAAATACTCT 540
QY 541 ATGTTTGCCAAATATTTGCTCTGCTGTAATGGGTAATAGATATATGCTCCCTTATGAC 600
DB 541 ATGTTTGCCAAATATTTGCTCTGCTGTAATGGGTAATAGATATATGCTCCCTTATGAC 600
QY 601 AAGAGACACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 AAGAGACACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 GCAAGTGAAGATCTCTATAGTAACGTGTAAGATGTAAGATGTAAGATGTAAGATGTAAG 720
DB 661 GCAAGTGAAGATCTCTATAGTAACGTGTAAGATGTAAGATGTAAGATGTAAGATGTAAG 720
QY 721 ATCAGTGTGCTGAGGACAGTTCGAAGGATCTGCAATTTGTTCAAAAGCATTTATTA 780
DB 721 ATCAGTGTGCTGAGGACAGTTCGAAGGATCTGCAATTTGTTCAAAAGCATTTATTA 780
QY 781 TACGATTTACACCGAGTACGCTATGTAGAGAGAGTACGATTAATCTACTAGATTG 840
DB 781 TACGATTTACACCGAGTACGCTATGTAGAGAGAGTACGATTAATCTACTAGATTG 840
QY 841 GCATCTACTAGATTATTAATATAGTCTGAGAAATTTGAGCATGTTTATATACACTT 900
DB 841 GCATCTACTAGATTATTAATATAGTCTGAGAAATTTGAGCATGTTTATATACACTT 900

```

```

QY 901 TCATTTCAAAATGTTTTCTCTGATCTGCTTTTATTTTACAAAAATATTTTCATACA 960
DB 901 TCATTTCAAAATGTTTTCTCTGATCTGCTTTTATTTTACAAAAATATTTTCATACA 960
QY 961 AATGGTAAAGAAACAAATCTATACAAACACTTTGGATTTTATATATAACTTTG 1020
DB 961 AATGGTAAAGAAACAAATCTATACAAACACTTTGGATTTTATATATAACTTTG 1020
QY 1021 TGATTTAAATTTACTGAATTTTATAGGCTGAAATTTTGAAGTGTATCTCATATGA 1080
DB 1021 TGATTTAAATTTACTGAATTTTATAGGCTGAAATTTTGAAGTGTATCTCATATGA 1080
QY 1081 CTAAGTCACTAAACCTGATGAAAGTAAATATATGTTCTGAGAACAAATGTACA 1140
DB 1081 CTAAGTCACTAAACCTGATGAAAGTAAATATATGTTCTGAGAACAAATGTACA 1140
QY 1141 AAAAGACAATTAATTTTACATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1141 AAAAGACAATTAATTTTACATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 CTAAGGCTAAGCATCTTCAAAAGCTTTTCCATATGCTGTCTTAATCTTTTCACTCAT 1260
DB 1201 CTAAGGCTAAGCATCTTCAAAAGCTTTTCCATATGCTGTCTTAATCTTTTCACTCAT 1260
QY 1261 TCACCTTCTTCCCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 1261 TCACCTTCTTCCCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1321 TAAACATGCTGATTTTATTTTCCCAAAATCATATCAATCCTTGAATTTTATATG 1380
DB 1321 TAAACATGCTGATTTTATTTTCCCAAAATCATATCAATCCTTGAATTTTATATG 1380
QY 1381 CATAAATTTTACAGTAATATGATCAAACTTCAATTTTAAACCTCTCTCTCTTGAC 1440
DB 1381 CATAAATTTTACAGTAATATGATCAAACTTCAATTTTAAACCTCTCTCTCTTGAC 1440
QY 1441 AAAATCTCTTAAAAAACAATATCAAGATATATAGTAAATACCTCCACTCAGACAGT 1500
DB 1441 AAAATCTCTTAAAAAACAATATCAAGATATATAGTAAATACCTCCACTCAGACAGT 1500
QY 1501 AGAAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560
DB 1501 AGAAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560
QY 1561 GGAGTATGGAAGGAAACATATGTAATTTACAGGGGAGAAAAATGACAAATGACGCTT 1620
DB 1561 GGAGTATGGAAGGAAACATATGTAATTTACAGGGGAGAAAAATGACAAATGACGCTT 1620
QY 1621 CACCAAGTATCAAAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1680
DB 1621 CACCAAGTATCAAAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1680
QY 1681 CTTTCTAAAAATTCATATCCCAATCTAATATGAGCTTAAACATCCGACAACTCAAGT 1740
DB 1681 CTTTCTAAAAATTCATATCCCAATCTAATATGAGCTTAAACATCCGACAACTCAAGT 1740
QY 1741 TGAAGGACATTTCTACAAAAATATCCCTGGGATTTTATAGATATCTCTCAAAAGCTT 1800
DB 1741 TGAAGGACATTTCTACAAAAATATCCCTGGGATTTTATAGATATCTCTCAAAAGCTT 1800
QY 1801 ATCAGTGAAGAAATGAAGGAAATCTGAGAAACATACAGACCAATGAGACTTAAGAGACA 1860
DB 1801 ATCAGTGAAGAAATGAAGGAAATCTGAGAAACATACAGACCAATGAGACTTAAGAGACA 1860
QY 1861 TGTAGGCAAAATGCAATGTGCTTCTTGATCAGATCCGGAACAGAAAAAGATCAGTAAT 1920
DB 1861 TGTAGGCAAAATGCAATGTGCTTCTTGATCAGATCCGGAACAGAAAAAGATCAGTAAT 1920
QY 1921 GAAAAAATGATGAAGTCTGAATGAATCTGAGATATTTTAAACAGTATGTTGATTTCT 1980
DB 1921 GAAAAAATGATGAAGTCTGAATGAATCTGAGATATTTTAAACAGTATGTTGATTTCT 1980
QY 1981 TAACTTGAACAAATATAGCAGGTAATGATAGATGATTAAGTGAAGAACTGAAGACTGG 2040

```

|||||
Db 1981 TATCTTGACAAATATAGAGGGTATGTAGTAACTGAGAACTGAACTGG 2040
OY 2041 GTGAGGCTATCTAGCAATCTCTCTACTATCTTACCAATTTTCGTAAGTGAAGAA 2100
|||||
Db 2041 GTGAGGCTATCTAGCAATCTCTCTACTATCTTACCAATTTTCGTAAGTGAAGAA 2100
OY 2101 GCATGCAAAATATAAGTATCTTGAAAAAATTTTAAAAA 2144
|||||
Db 2101 GCATGCAAAATATAAGTATCTTGAAAAAATTTTAAAAA 2144
RESULT 2
HSPG28 2138 bp mRNA linear PRI 04-MAR-1996
LOCUS H.sapiens mRNA for SGP28 protein.
DEFINITION X94323
ACCESSION X94323
VERSION X94323.1 GI:1213612
KEYWORDS glycoprotein; SGP28 protein.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2138)
AUTHORS Kjeldsen, L., Cowland, J.B., Johnsen, A.H. and Borregaard, N.
TITLE SGP28, a novel matrix glycoprotein in specific granules of human
neutrophils with similarity to a human testis-specific gene product
and a rodent sperm-coating glycoprotein
JOURNAL FEBS Lett. 380 (3), 246-250 (1996)
MEDLINE 96186934
PUBMED 8601434
REFERENCE 2 (bases 1 to 2138)
AUTHORS Cowland, J.B.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1995) J.B. Cowland, Granulocyte Research
Laboratory, Dept of Hematology, National Univ. Hosp.,
Rigshospitalet L-4041, 9 Blegdamsvej, 2100 Copenhagen, DENMARK
FEATURES
source
1. 2138
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="bone marrow"
41..778
/codon_start=1
/product="SGP28 protein"
/protein_id="CAA63984.1"
/db_xref="GI:1213612"
/db_xref="SWISS-PROT:P54108"
/translation="MTLFFVLFLVAGLLSPFPAEDKDPATFALLITQYQVREIVN
KHNLRRRAVSPPARMLKMKEMKKAANAKNANCNYHSPKDMTSLCKGKGNLYM
SSAPSSWSQAIOSMFDEYNDPFGVGPKNPAVHYTQVYVSSLYVCGNAYCPNO
KYLKYYVCOYCPAGNMANRLVPEQGAFCAPCAVCPNCDGDLCTNCKKEDLYSNCKS
IKLTLTCKQLVDRSCAKSCNCSNSIY"
BASE COUNT 728 a 404 c 383 g 623 t
ORIGIN
Query Match 99.6%; Score 2134.8; DB 9; Length 2138;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

|||||
Db 181 TGAAGTGAAGAGAGAGATATCTCCCTGCGCAAAACATGCTGAAGTGAAGAACAA 240
OY 242 AGAGGCTGCGCAAAATGCGCCAAAAGTGGGCAACACAGTGCATTAAGACACAGTAACC 301
|||||
Db 241 AGAGGCTGCGCAAAATGCGCCAAAAGTGGGCAACACAGTGCATTAAGACACAGTAACC 300
OY 302 AAAGATGCAATGACAAATGCTAAATGTGTGAGAAATCTCTACATGTCAAGTCCCCAG 361
|||||
Db 301 AAAGATGCAATGACAAATGCTAAATGTGTGAGAAATCTCTACATGTCAAGTCCCCAG 360
OY 362 CTGATGCTCACAAGCAATCCAAAGCTGTTTGTATGAGTACATGATTTGACTTGTGTG 421
|||||
Db 361 CTGATGCTCACAAGCAATCCAAAGCTGTTTGTATGAGTACATGATTTGACTTGTGTG 420
OY 422 AGGGCCAAAGACTGCCAAGCAGTGGTGGACATTTATACAGAGTTGTTGGTACTCTTC 481
|||||
Db 421 AGGGCCAAAGACTGCCAAGCAGTGGTGGACATTTATACAGAGTTGTTGGTACTCTTC 480
OY 482 ATACCTGTTGATGTGAAATGCTACTGTGCCAATCAAAAAGTTCTAAATACTACTA 541
|||||
Db 481 ATACCTGTTGATGTGAAATGCTACTGTGCCAATCAAAAAGTTCTAAATACTACTA 540
OY 542 TGTTCGCCAATATTTGCTCTGCTGTAATGGCTAATAGCTATATGTCCCTATGAACA 601
|||||
Db 541 TGTTCGCCAATATTTGCTCTGCTGTAATGGCTAATAGCTATATGTCCCTATGAACA 600
OY 602 AGGAGCACTTGTGCCAGTGGCCAGATACCTGAGATGAGATGAGATGAGCAATGGTTG 661
|||||
Db 601 AGGAGCACTTGTGCCAGTGGCCAGATACCTGAGATGAGATGAGATGAGCAATGGTTG 660
OY 662 CAAATGAGAGATCTGTATAGTAACTGTAAAGCTTGAACCTCAATTAACCTGTAACA 721
|||||
Db 661 CAAATGAGAGATCTGTATAGTAACTGTAAAGCTTGAACCTCAATTAACCTGTAACA 720
OY 722 TCACTGTTGATGAGGACAGTTGCAAGGCACTCTGCAATGTTCAAACAGCATTTTAAAT 781
|||||
Db 721 TCACTGTTGATGAGGACAGTTGCAAGGCACTCTGCAATGTTCAAACAGCATTTTAAAT 780
OY 782 ACCGATTACACACGAGTAGGAGGCTATGTAGAGAGAGTCAATATCTACTTATGATTGG 841
|||||
Db 781 ACCGATTACACACGAGTAGGAGGCTATGTAGAGAGAGTCAATATCTACTTATGATTGG 840
OY 842 CATCTAGTATGATTAACATATCTAGCTGAGAAATGTAGGAGATGTTGATACACATTT 901
|||||
Db 841 CATCTAGTATGATTAACATATCTAGCTGAGAAATGTAGGAGATGTTGATACACATTT 900
OY 902 GATTTCAATGTTTTCTTCTGATCTGCTTTTATTTTCAAAAATATTTTCTATACAA 961
|||||
Db 901 GATTTCAATGTTTTCTTCTGATCTGCTTTTATTTTCAAAAATATTTTCTATACAA 960
OY 962 ATGTTTAAAGAAACAAATCTATAACAACAACACTTTGATTTTATATATAAATTTGT 1021
|||||
Db 961 ATGTTTAAAGAAACAAATCTATAACAACAACAACACTTTGATTTTATATATAAATTTGT 1020
OY 1022 GATTTAAATTTACTGAATTTAATTTAGGCTGAATAATTTTGAAGTTGATTTCTATATGAC 1081
|||||
Db 1021 GATTTAAATTTACTGAATTTAATTTAGGCTGAATAATTTTGAAGTTGATTTCTATATGAC 1080
OY 1082 TAAGTTCACTAAACCTGATGAAAGTGAATAATTTGTTCTAGAACAAATGTACAA 1141
|||||
Db 1081 TAAGTTCACTAAACCTGATGAAAGTGAATAATTTGTTCTAGAACAAATGTACAA 1140
OY 1142 AAAGAACATATATATTTTCAATGAACCTTGGCTGTAGTTGGCTTCTAGTCCACTC 1201
|||||
Db 1141 AAAGAACATATATATTTTCAATGAACCTTGGCTGTAGTTGGCTTCTAGTCCACTC 1200
OY 1202 TAAAGCTAAGCATCTTCAAAAGAGTTTCCCATATGCTGTCTTAATTTCTTTTACATCAT 1261
|||||
Db 1201 TAAAGCTAAGCATCTTCAAAAGAGTTTCCCATATGCTGTCTTAATTTCTTTTACATCAT 1260
OY 1262 CACCTTCTTCCCAATCATGTGCTGGCATCTGCACAAATGAGTGAAGTGTTCCTCCT 1321
|||||

Db 1261 CACCCTTTCCCAATCATCTGGCTGGCATCCCTACAAATGAGTTGAAGCTGTCTCCT 1320
Oy 1322 AAAACAATCTGCACTTTTATTTTGGCAAAATACATACATCTTTGAATTTTATTCGTC 1381
Db 1321 AAAACAATCTGCACTTTTATTTTGGCAAAATACATACATCTTTGAATTTTATTCGTC 1380
Oy 1382 ATAAATTTTACAGTAAGTAATGATCAAACTTCATTTTAAACCTCTCTCTCTTTGACA 1441
Db 1381 ATAAATTTTACAGTAAGTAATGATCAAACTTCATTTTAAACCTCTCTCTCTTTGACA 1440
Oy 1442 AAACCTCTTAAAGTAAGTAATGATCAAACTTCATTTTAAACCTCTCTCTCTTTGACA 1501
Db 1441 AAACCTCTTAAAGTAAGTAATGATCAAACTTCATTTTAAACCTCTCTCTCTTTGACA 1500
Oy 1502 GAACTCAGTCT 1561
Db 1501 GAACTCAGTCT 1560
Oy 1562 GAGTATGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGG 1621
Db 1561 GAGTATGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGG 1620
Oy 1622 ACCAAGTATCAAAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1681
Db 1621 ACCAAGTATCAAAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Oy 1682 TTTCTAAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1741
Db 1681 TTTCTAAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1740
Oy 1742 GAAAGACATTTCTACAAATATCTCTGGGGATTTTGAAGTATCTCTCAAACTGTAAAAA 1801
Db 1741 GAAAGACATTTCTACAAATATCTCTGGGGATTTTGAAGTATCTCTCAAACTGTAAAAA 1800
Oy 1802 TCATGGAATTAAGGGAATCTCTGAGAAACATCAGAGCAATGAGACTAAGGAGACAT 1861
Db 1801 TCATGGAATTAAGGGAATCTCTGAGAAACATCAGAGCAATGAGACTAAGGAGACAT 1860
Oy 1862 GTGAGCCAAATGCAATGTGCTTCTTGATCAGATCCTGGAAACAGAAAAAGATCAGTAATG 1921
Db 1861 GTGAGCCAAATGCAATGTGCTTCTTGATCAGATCCTGGAAACAGAAAAAGATCAGTAATG 1920
Oy 1922 AAAAAGCGATGAAGTCTGGAATAGAAATCTGGAATTTTAACTAGTGTGATTTCTT 1981
Db 1921 AAAAAGCGATGAAGTCTGGAATAGAAATCTGGAATTTTAACTAGTGTGATTTCTT 1980
Oy 1982 AATCTTGAACAATATAGCAGGGAATGTAAGATGATAAGTGTAGAGAACTGAACCTGGG 2041
Db 1981 AATCTTGAACAATATAGCAGGGAATGTAAGATGATAAGTGTAGAGAACTGAACCTGGG 2040
Oy 2042 TGAGGGCTATCTAGGAATTTCTGTACTATCTTACCAAAATTTTGGGTAAGTCTAAGAA 2101
Db 2041 TGAGGGCTATCTAGGAATTTCTGTACTATCTTACCAAAATTTTGGGTAAGTCTAAGAA 2100
Oy 2102 CAATGCAAAATTAAGGATCTTGAAGAAAAA 2139
Db 2101 CAATGCAAAATTAAGGATCTTGAAGAAAAA 2138

RESULT 3
AX335634
LOCUS AX335634 2128 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 6143 from Patent WO0194629.
ACCESSION AX335634
VERSION AX335634.1 GI:18126353
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Young, P. E., Augustus, M., Carter, K. C., Ehner, R., Endress, G.,
Horrikan, S., Soppet, D. R. and Weaver, Z.

TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL gene sets
Patent: WO 0194629-A 6143 13-DEC-2001;
Avallon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1. 2128
/organism="Homo sapiens"
/db_xref="Laxon:9606"
BASE COUNT 734 a 397 c 380 g 617 t
ORIGIN
Query Match 98.6%; Score 2113.2; DB 6; Length 2128;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2115; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 27 CTGGAAGCCACTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 86
Db 1 CTGGAAGCCACTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Oy 87 CCATCTTTCCAGCAAAATGAAGTAAGATCCCGCTTTTACTGCTTTGTAAACCAACCA 146
Db 61 CCATCTTTCCAGCAAAATGAAGTAAGATCCCGCTTTTACTGCTTTGTAAACCAACCA 120
Oy 147 ACACAAGTGAAGGAGAGATTTGTAAATGAACCAATGAAGTGAAGAGAGAGAGATCTCC 206
Db 121 ACACAAGTGAAGGAGAGATTTGTAAATGAACCAATGAAGTGAAGAGAGAGAGATCTCC 180
Oy 207 CTGCGCAGAACATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 266
Db 181 CTGCGCAGAACATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Oy 267 TGGGCAACCACTGCAATTTACAGACAGTAACCCAAAGATGATGATGATGATGATGATGAT 326
Db 241 TGGGCAACCACTGCAATTTACAGACAGTAACCCAAAGATGATGATGATGATGATGATGAT 300
Oy 327 TGTGTGAGAAATCTTACATGTCAGAGTGGCCCACTGATGTCACAGCAATCCAAAGC 386
Db 301 TGTGTGAGAAATCTTACATGTCAGAGTGGCCCACTGATGTCACAGCAATCCAAAGC 360
Oy 387 TGTGTGAGAAATCTTACATGTCAGAGTGGCCCACTGATGTCACAGCAATCCAAAGC 446
Db 361 TGTGTGAGAAATCTTACATGTCAGAGTGGCCCACTGATGTCACAGCAATCCAAAGC 420
Oy 447 GTTGACATTAATACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 506
Db 421 GTTGACATTAATACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Oy 507 TACTGTCCCAATCAAAAGTCTAAATTAATCTAATGTTGGCAATATGTCCTGCTGCT 566
Db 481 TACTGTCCCAATCAAAAGTCTAAATTAATCTAATGTTGGCAATATGTCCTGCTGCTGCT 540
Oy 567 AATTGGGTAATAGACTATATGTCCTTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626
Db 541 AATTGGGTAATAGACTATATGTCCTTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Oy 627 GATAAGTGTGAG 686
Db 601 GATAAGTGTGAG 660
Oy 687 TGTAAAGTTGAAAGTGTGAAAGTGTGAAAGTGTGAAAGTGTGAAAGTGTGAAAGTGTGAA 746
Db 661 TGTAAAGTTGAAAGTGTGAAAGTGTGAAAGTGTGAAAGTGTGAAAGTGTGAAAGTGTGAA 720
Oy 747 GCATCTGCAATTTGTCAACAGATTTATTAATAGCAGATTAACAGAGAGAGAGAGAGAG 806
Db 721 GCATCTGCAATTTGTCAACAGATTTATTAATAGCAGATTAACAGAGAGAGAGAGAGAG 780
Oy 807 TGTAG 866
Db 781 TGTAG 840
Oy 867 AGCTAG 926

[illegible]

OY	1227	TTTCATATGCTGTCTTAATTTCTTTTACATCATTCACCTCTTCCCATCATCTGGCT	1286
Db	1201	TTTCCCAATGCTGTCTTAATTTCTTTTACATCATTCACCTCTTCCCAATCATCTGGCT	1260
OY	1287	GGCATCCTCACAATTTGATTTGACCTGTCTCTCTTAACAAATCCGACTTTTATTTTGC	1346
Db	1261	GGCATCCTCACAATTTGATTTGACCTGTCTCTCTTAACAAATCCGACTTTTATTTTGC	1320
OY	1347	CAAAAATCAATCAATCCTTTGAATTTTTTATCTGCATTAATTTTACAGTGAATATGATC	1406
Db	1321	CAAAATCAATCAATCCTTTGAATTTTTTATCTGCATTAATTTTACAGTGAATATGATC	1380
OY	1407	AAACCTTATTTTAAACCTCTCTTCTTTGACAAAATCTTCTTTAAAAAAGATACAAAG	1466
Db	1381	AAACCTTATTTTAAACCTCTCTTCTTTGACAAAATCTTCTTTAAAAAAGATACAAAG	1440
OY	1467	ATAATATAGTAAATATACCTCTCAGTCAAGGAGTGAAGAACTCAGTCCCTCTGTGAGT	1526
Db	1441	ATAATATAGTAAATATACCTCTCAGTCAAGGAGTGAAGAACTCAGTCCCTCTGTGAGT	1500
OY	1527	CTTCACTAAAATCAAGTCACTCACTTCCAAAGAGTGAATATGAAAAGGAACATAGTAA	1586
Db	1501	CTTCACTAAAATCAAGTCACTCACTTCCAAAGAGTGAATATGAAAAGGAACATAGTAA	1560
OY	1587	CTTTACAGGGAGAAAATGACAAATGACGCTTTCCACCAAGTATCAAAATTAGCTCAC	1646
Db	1561	CTTTACAGGGAGAAAATGACAAATGACGCTTTCCACCAAGTATCAAAATTAGCTCAC	1620
OY	1647	CAGTATTAAGTCAATTCAGATTTGTTGTAGATAATCTTTCTAAAAATTCATATCCCAATC	1706
Db	1621	CAGTATTAAGTCAATTCAGATTTGTTGTAGATAATCTTTCTAAAAATTCATATCCCAATC	1680
OY	1707	TAAATTATAGCTAAACATCCAGCAAACTCAAGTTGAAGGACATTTCTACAAAATATCCCT	1766
Db	1681	TAAATTATAGCTAAACATCCAGCAAACTCAAGTTGAAGGACATTTCTACAAAATATCCCT	1740
OY	1767	GGGGATTTTGTAGATATTTCCCTCAAAACGTAAATATGATGAAAATTAAGGAATCCCTGAG	1826
Db	1741	GGGGATTTTGTAGATATTTCCCTCAAAACGTAAATATGATGAAAATTAAGGAATCCCTGAG	1800
OY	1827	AAACATCACAGACACATGAGACTAAGAGAGACATGTGAGCCAAATTCATATGCTTCTT	1886
Db	1801	AAACATCACAGACACATGAGACTAAGAGAGACATGTGAGCCAAATTCATATGCTTCTT	1860
OY	1887	GGATCAGATCCTGGACAGAAAAGATCAGTAATGAAAACCTGATGAAGTCTGAATAGA	1946
Db	1861	GGATCAGATCCTGGACAGAAAAGATCAGTAATGAAAACCTGATGAAGTCTGAATAGA	1920
OY	1947	ATTCGAGATTTTTTAAACAGTACGTGATTTCTTATCTTGACAAATATAGAGAGGTAA	2006
Db	1921	ATTCGAGATTTTTTAAACAGTACGTGATTTCTTATCTTGACAAATATAGAGAGGTAA	1980
OY	2007	TGTAAGATGATTAAGTAAAGAACTGAGTGGTGAAGGCTATCTAGAGAAATTCCTGT	2066
Db	1981	TGTAAGATGATTAAGTAAAGAACTGAGTGGTGAAGGCTATCTAGAGAAATTCCTGT	2040
OY	2067	ACTATCTTACCAAAATTTTGGTAAAGCTTAAGTAAAGAAAGCAATCCAAATTAAGATATCTTGA	2126
Db	2041	ACTATCTTACCAAAATTTTGGTAAAGCTTAAGTAAAGAAAGCAATCCAAATTAAGATATCTTGA	2100
OY	2127	AAAAAAAAAAAAAAAAAAAAA 2144	
Db	2101	AAAAAAAAAAAAAAAAAAAAA 2118	

QY 1072 CTCATATGACTAAGTTCACATAAAACCTGATGTGAAAGTGAATATGTTCTAGAACAA 1131
 |||||||
 Db 79686 CTCATATGACTAAGTTCACATAAAACCTGATGTGAAAGTGAATATGTTCTAGAACAA 79627
 |||||||
 QY 1132 AATGTACAAAAGACAAATATATTTTACATGAGACCTTGCTGTAGTTGCTTCTCT 1191
 |||||||
 Db 79626 AATGTACAAAAGACAAATATATTTTACATGAGACCTTGCTGTAGTTGCTTCTCT 79567
 |||||||
 QY 1192 AGCTCCACTAAGCTTAAGATCTTCAAAAGCTTTCCATATGCTGTATATTTCT 1251
 |||||||
 Db 79566 AGCTCCACTAAGCTTAAGATCTTCAAAAGCTTTCCATATGCTGTATATTTCT 79567
 |||||||
 QY 1252 TTCACATATGACCTTCTTCCCAATATCTGCTGGCATCTCCCAATGATGTGAAGC 1311
 |||||||
 Db 79506 TTCACATATGACCTTCTTCCCAATATCTGCTGGCATCTCCCAATGATGTGAAGC 79447
 |||||||
 QY 1312 TGTTCCTCTTAAACATCTGACTTTTATTTTGGCAAAATCATACATCTTGAATT 1371
 |||||||
 Db 79446 TGTTCCTCTTAAACATCTGACTTTTATTTTGGCAAAATCATACATCTTGAATT 79387
 |||||||
 QY 1372 TTTATCTGCAATAATTTTACAGATATGATCAAACTTCATTTTAAACCTCTCT 1431
 |||||||
 Db 79386 TTTATCTGCAATAATTTTACAGATATGATCAAACTTCATTTTAAACCTCTCT 79327
 |||||||
 QY 1432 CTCTTTGACAAAACCTCTCTTAAAAAGAAATCAAGATATATAGTAAATACCTCTCACT 1491
 |||||||
 Db 79326 CTCTTTGACAAAACCTCTCTTAAAAAGAAATCAAGATATATAGTAAATACCTCTCACT 79267
 |||||||
 QY 1492 CAAGGAGTAACTACATCTCTCTCTGAGCTCTCAATAAATCAGAGCTCACT 1551
 |||||||
 Db 79266 CAAGGAGTAACTACATCTCTCTCTGAGCTCTCAATAAATCAGAGCTCACT 79207
 |||||||
 QY 1552 CCAAGAGTGAATGGAAGGGAAGGGAACATAGTAATTTACAGGGGAGAAAATGACAAA 1611
 |||||||
 Db 79206 CCAAGAGTGAATGGAAGGGAAGGGAACATAGTAATTTACAGGGGAGAAAATGACAAA 79147
 |||||||
 QY 1612 TGACGCTCTTCAACAGATATCAAAATTAAGCTCACCAGTATAGTATTCAGATTGTT 1671
 |||||||
 Db 79146 TGACGCTCTTCAACAGATATCAAAATTAAGCTCACCAGTATAGTATTCAGATTGTT 79087
 |||||||
 QY 1672 CTAGATATCTTTCTTAAAAATTCATATCCCAATCTATATATGAGCTAAACATCCACAA 1731
 |||||||
 Db 79086 CTAGATATCTTTCTTAAAAATTCATATCCCAATCTATATGAGCTAAACATCCACAA 79027
 |||||||
 QY 1732 AACTCAAGTGAAGACATTTCTACAAATATCCCTGGGGTATTTAGAGTATTCCTCAAA 1791
 |||||||
 Db 79026 AACTCAAGTGAAGACATTTCTACAAATATCCCTGGGGTATTTAGAGTATTCCTCAAA 78967
 |||||||
 QY 1792 ACTGTAAAAATCATGAGAAATAGGGAATCCTGAGAAAACATCACAGACCATGAGACT 1851
 |||||||
 Db 78966 ACTGTAAAAATCATGAGAAATAGGGAATCCTGAGAAAACATCACAGACCATGAGACT 78907
 |||||||
 QY 1852 AAGGAGATGTGAGCCAAATGCAATGTGCTCTGATGCAGATCCTGGAACAGAAAAG 1911
 |||||||
 Db 78906 AAGGAGATGTGAGCCAAATGCAATGTGCTCTGATGCAGATCCTGGAACAGAAAAG 78847
 |||||||
 QY 1912 ATCAATAATGAAAAAATGATGAGTCTGAATAGATCTGAGATTTTAAACAGTAGTG 1971
 |||||||
 Db 78846 ATCAATAATGAAAAAATGATGAGTCTGAATAGATCTGAGATTTTAAACAGTAGTG 78787
 |||||||
 QY 1972 TTGATTTTATATCTTGCAAAATATAGCAGGGTAAATGTAAGATGTAACGTTAGAGAAC 2031
 |||||||
 Db 78786 TTGATTTTATATCTTGCAAAATATAGCAGGGTAAATGTAAGATGTAACGTTAGAGAAC 78727
 |||||||
 QY 2032 TGAACCTGGGTGAGGCGATCTAGGAATCTCTGACTATCTTCAAAATTTTGGTAAAG 2091
 |||||||
 Db 78726 TGAACCTGGGTGAGGCGATCTAGGAATCTCTGACTATCTTCAAAATTTTGGTAAAG 78667
 |||||||
 QY 2092 TCTAAGAAGCAATGCAAAATTAAGATATCTTGAAGAAAAAATTAAGAA 2141
 |||||||
 Db 78666 TCTAAGAAGCAATGCAAAATTAAGATATCTTGAAGAAAAAATTAAGAA 78617
 |||||||

RESULT 6 HSDJ442L6/c LOCUS DEFINITION

HSDJ442L6 118524 bp DNA linear PRI 30-SEP-2000
 Human DNA sequence from clone RPJ-442L6 on chromosome 6. Contains
 the RPX1 gene encoding a Testis specific protein 1 (CRISP-2,
 GAPD5), the 5' end of RHAG gene encoding Rhesus blood
 group-associated glycoprotein (RH50a), the 3' end of SPG28
 encoding a cysteine-rich secretory protein-3 (CRISP-3), ESTs, STSS
 and GSSs, complete sequence.

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

AL121950.8 GI:9581779
 HTG.
 human.
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE AUTHORS TITLE JOURNAL

Parker, A.
 Direct Submission
 Submitted (29-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk
 On Jul 28, 2000 this sequence version replaced gi:9408740.

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence has been finished according to sequence map criteria
 as follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known
 annotated human repeat sequence elements (e.g. Alu). Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key.

The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em., EMBL; SW., SWISSPROT; Tr., TREMBL; Wp., WormPeP; Information
 on the WormPeP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 This sequence was generated from part of bacterial clone/corliss of human
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
 group. Further information can be found at
 http://www.sanger.ac.uk/Chp6
 RPJ-442L6 is from the library RPJ-3 constructed at the Rosewell
 Park Cancer Institute by the group of Pleier de Jong. For further
 details see http://bacpac.med.buffalo.edu/

VECTOR: pCIPAC2
 This sequence is the entire insert of clone RPJ-442L6 The true left
 end of clone RPJ-417L20 is at 85775 in this sequence.
 Location/Qualifiers

FEATURES

source

1. 118524

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="6"

/clone="RPJ-442L6"

/clone_id="RPJ-3"

complement(1..25721)

/gene="RHAG"

complement(join<1..27,1279..1416,3569..3735,4506..4653,
6950..7100,8061..8244,25538..25721))

/product="dj442L6.1 (Rhesus blood group-associated
glycoprotein (RH50A))"

/gene="RHAG"

Em:AF187847 Em:XM6594 Em:AF179682 Em:AF031548

Em:AF17847 Em:AF17844 Em:AF058917 Em:AF179684

Em:AB015467

/evidence="not_experimental"

complement(join<1..27,1279..1416,3569..3735,4506..4653,
6950..7100,8061..8244,25538..25694))

/gene="RHAG"

/note="Continues in clone dj417L20 (Em:AL121974)"

CDS


```
match: proteins: Tr:09V3T3 Tr:088300 Tr:07484 Tr:09UBD6
Tr:09QXP0"
/codon_start=1
/evidence=not_experimental
/product="J442L6.1 (Rhesus blood group-associated
glycoprotein (RH50A))"
/protein_id="CAC10519.1"
/db_xref="GI:10443472"
/translation="MRTFPLMAIVLEIAMIVLGLPEYETDQVLEQLNITKPTD
GIFELPLFDQVHMTFVGFPLMTFLKATGSSVGINLVAALGLQMTIVQCILO
SOGKFNIGIKMINADFSAAATVLISFGAVGKTSPTQMLMTLIEVFANHELV
EIRFASDAGSMTHAFGAYGLAVAGILYSGRKGNESAVSDLPAMIGLFL
MMPSPNSAIEPDKKOCRAIVNTYPSLACVLTAFFSSIVEHGRKLNHVNIONAT
LAGVAVGTCAADMAIHPRGSMIISAGMSVLGKFLPTTKLR"
repeat_region
257..446
/notes="LIM7 repeat: matches 5986..6165 of consensus"
623..930
/notes="AluX repeat: matches 1..301 of consensus"
1482..1913
/notes="match: GSS: Em:AQ813163"
1952..2249
/notes="AluX repeat: matches 3..300 of consensus"
2602..2746
/notes="LIM3 repeat: matches 7582..7740 of consensus"
2785..2832
/notes="LIM3 repeat: matches 7478..7525 of consensus"
3064..3383
/notes="AluSg repeat: matches 1..311 of consensus"
3823..4131
/notes="AluSg repeat: matches 1..305 of consensus"
4288..4406
/notes="FLAM_A repeat: matches 1..122 of consensus"
4820..5005
/notes="MIR repeat: matches 20..211 of consensus"
5084..5257
/notes="FRAM repeat: matches 0..172 of consensus"
5951..6387
/notes="L2 repeat: matches 2246..2698 of consensus"
6673..6828
/notes="MIR repeat: matches 95..262 of consensus"
complement(7407..8012)
/genes="RHAG"
/notes="match: GSS: Em:B87311"
7685..8049
/notes="MTR1A1 repeat: matches 1..365 of consensus"
8340..8442
/notes="MIR repeat: matches 4..106 of consensus"
9022..9137
/notes="MIR repeat: matches 49..172 of consensus"
9432..9712
/notes="AluX repeat: matches 1..262 of consensus"
10074..10388
/notes="LTR16A repeat: matches 14..381 of consensus"
10815..10880
/notes="33 copies 2 mer tt 66% conserved"
10939..11098
/notes="THE1B repeat: matches 1..167 of consensus"
11099..11389
/notes="AluSg repeat: matches 2..297 of consensus"
11390..11609
/notes="THE1B repeat: matches 167..364 of consensus"
11713..12019
/notes="AluY repeat: matches 1..304 of consensus"
12444..12662
/notes="MER20 repeat: matches 1..217 of consensus"
13096..13385
/notes="AluX repeat: matches 2..290 of consensus"
13474..13553
/notes="MIR repeat: matches 74..182 of consensus"
complement(13493..13920)
/genes="RHAG"
/notes="match: GSS: Em:AQ813968"
13541..13606
repeat_region

misc_feature
13971..14526
/notes="L2 repeat: matches 2637..2707 of consensus"
14139..14391
/notes="match: GSS: Em:AQ539951"
14620..14983
/notes="MIR repeat: matches 5..262 of consensus"
14984..16480
/notes="THE1B repeat: matches 1..364 of consensus"
/notes="THE1B-INTERNAL repeat: matches 1..1580 of
consensus"
16481..16840
/notes="THE1B repeat: matches 1..364 of consensus"
17057..17296
/notes="AluY repeat: matches 31..282 of consensus"
17297..18994
/notes="PFR5 repeat: matches 743..1686 of consensus"
19639..19831
/notes="LIM3 repeat: matches 7448..7634 of consensus"
19863..20630
/notes="LIM4 repeat: matches 6630..7474 of consensus"
20891..21016
/notes="Charlie1a repeat: matches 1..128 of consensus"
21013..21609
/notes="Charlie1a repeat: matches 331..933 of consensus"
22116..22212
/notes="L2 repeat: matches 2589..2695 of consensus"
22790..22844
/notes="L2 repeat: matches 2617..2671 of consensus"
22860..23218
/notes="LIM4 repeat: matches 5950..6299 of consensus"
23216..23511
/notes="LIM2 repeat: matches 5508..5806 of consensus"
23683..23832
/notes="LIM3 repeat: matches 6152..6302 of consensus"
23839..23880
/notes="21 copies 2 mer tg 88% conserved"
24216..24538
/notes="match: GSS: Em:AQ071395"
27538..27945
/notes="L2 repeat: matches 2356..2750 of consensus"
28904..29192
/notes="AluX repeat: matches 1..289 of consensus"
29656..30138
/notes="Charlie4a repeat: matches 1..507 of consensus"
30286..30707
/notes="match: GSS: Em:AQ465695"
30451..30583
/notes="MIR repeat: matches 1..134 of consensus"
30513..30810
/notes="match: GSS: Em:AQ114043"
30770..31070
/notes="AluY repeat: matches 1..303 of consensus"
31388..31459
/notes="36 copies 2 mer ta 69% conserved"
31808..32127
/notes="AluSg repeat: matches 1..309 of consensus"
32391..32514
repeat_region

Query Match 69.0%; Score 1480.4; DB 9; Length 118524;
Best Local Similarity 99.6%; Pred. No. 4e-263;
Matches 1484; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 652 CCAATGCTGCAGACGATCTATAGTAACTAAAGTTTAACTACATTA 711
|||||
DB 117739 CCAATGCTGCAGATAGAGATCTATAGTAAAGTTTAACTACATTA 117680

OY 712 CCTGTAACATCAGTGTGAGGACAGTGCAGGACCTCCTGCAATGTTCAACAGCA 771
|||||
DB 117679 CCTGTAACATCAGTGTGAGGACAGTGCAGGACCTCCTGCAATGTTCAACAGCA 117620

OY 772 TTTATTAATACGATTAACACACGAGTACGAGGCTATAGAGAGATCAATATCTAC 831
|||||
DB 117619 TTTATTAATACGATTAACACACGAGTACGAGGCTATAGAGAGATCAATATCTAC 117560
```

Oy 832 TTAGATTGGCATCTACCTAGATTATACATATAGTGAAGAAATTTAGGATGTTG 891
 |||||
 Db 117559 TTAGATTGGCATCTACCTAGATTATACATATAGTGAAGAAATTTAGGATGTTG 117500
 Oy 892 ATACACATTGATTTCAAAATGTTTTCTCTGATCTGCTTTTATTTTACAAAATAT 951
 |||||
 Db 117499 ATACACATTGATTTCAAAATGTTTTCTCTGATCTGCTTTTATTTTACAAAATAT 117440
 Oy 952 TTTTCATCAAAATGTTTTTAAAGAACAAATCTATACAAACCTTGGATTTTATATA 1011
 |||||
 Db 117439 TTTTCATCAAAATGTTTTTAAAGAACAAATCTATACAAACCTTGGATTTTATATA 117380
 Oy 1012 TAAACTTTGATTTTAAATTTACTGAAATTTAATTTAGGGGAAATTTTAAAGTTGAT 1071
 |||||
 Db 117379 TAAACTTTGATTTTAAATTTACTGAAATTTAATTTAGGGGAAATTTTAAAGTTGAT 117320
 Oy 1072 CTCATATGACATAGTTACATAAACCCCTGATTTGAAGTGAATAATATGTTCTTAGACA 1131
 |||||
 Db 117319 CTCATATGACATAGTTACATAAACCCCTGATTTGAAGTGAATAATATGTTCTTAGACA 117260
 Oy 1132 AATGTACAAAAGAACATATATTTTACATGAAACCTTGGCTGTAGTGGCTTCT 1191
 |||||
 Db 117259 AATGTACAAAAGAACATATATTTTACATGAAACCTTGGCTGTAGTGGCTTCT 117200
 Oy 1192 ACCCTCACTCTAAGGCTAAGCATCTTCAAGAGAGTTTCCCATATGCTGCTTATTTCT 1251
 |||||
 Db 117199 ACCCTCACTCTAAGGCTAAGCATCTTCAAGAGAGTTTCCCATATGCTGCTTATTTCT 117140
 Oy 1252 TTCATCATCTACCCCTCTTCCCAATCATGTGGCTGGACCTCTACAAATTTGAGTTGAGC 1311
 |||||
 Db 117139 TTCATCATCTACCCCTCTTCCCAATCATGTGGCTGGACCTCTACAAATTTGAGTTGAGC 117080
 Oy 1312 TGTCTCTCTTAAACAAATCTGACTTTATTTTGGCAAAATCAATATCAATCTTTGAAT 1371
 |||||
 Db 117079 TGTCTCTCTTAAACAAATCTGACTTTATTTTGGCAAAATCAATATCAATCTTTGAAT 117020
 Oy 1372 TTTTATCTGCATTAATTTTACAGTAATATGATCAAAACCTTCAATTTTAAACCTCTCT 1431
 |||||
 Db 117019 TTTTATCTGCATTAATTTTACAGTAATATGATCAAAACCTTCAATTTTAAACCTCTCT 116960
 Oy 1432 CTCCTTGACAAAATCTCTCTTAAAAAAGAAATACAGATATATAGTAATATACCTCCACT 1491
 |||||
 Db 116959 CTCCTTGACAAAATCTCTCTTAAAAAAGAAATACAGATATATAGTAATATACCTCCACT 116900
 Oy 1492 CAAGGAGTAGAAGTCACTCCCTCTCTCTGTAGTCTTCACTTAAATCAGTACTCATT 1551
 |||||
 Db 116899 CAAGGAGTAGAAGTCACTCCCTCTCTCTGTAGTCTTCACTTAAATCAGTACTCATT 116840
 Oy 1552 CCAGAGATGAGATAGGAAAGAAACATATAGTAACTTACAGGGGAGAAATATGACAA 1611
 |||||
 Db 116839 CCAGAGATGAGATAGGAAAGAAACATATAGTAACTTACAGGGGAGAAATATGACAA 116780
 Oy 1612 TGACCTCTTCAACCAAGTATCAAAATTAACGTCACAGTATAGTCAATTCAGATTGTT 1671
 |||||
 Db 116779 TGACCTCTTCAACCAAGTATCAAAATTAACGTCACAGTATAGTCAATTCAGATTGTT 116720
 Oy 1672 CTAGTATCTTTCTTAAATTTCAATATCCCATATTAATTTAGCTTAAACATCCAGCA 1731
 |||||
 Db 116719 CTAGTATCTTTCTTAAATTTCAATATCCCATATTAATTTAGCTTAAACATCCAGCA 116660
 Oy 1732 AACTCAATTTGAGAGATCTTCAAAATATCCCTGGGGATTTTAAAGTATTTCTCAAA 1791
 |||||
 Db 116659 AACTCAATTTGAGAGATCTTCAAAATATCCCTGGGGATTTTAAAGTATTTCTCAAA 116600
 Oy 1792 ACTGTAAATATCAGGAAATTAAGGAATCTGAGAAACCAATCAGACACATAGACT 1851
 |||||
 Db 116599 ACTGTAAATATCAGGAAATTAAGGAATCTGAGAAACCAATCAGACACATAGACT 116540
 Oy 1852 AAGGAGATGTGAGCCAAATGCAATGTGCTTTCTTGATCAGATCTTGGACAGAGAAAAG 1911
 |||||
 Db 116539 AAGGAGATGTGAGCCAAATGCAATGTGCTTTCTTGATCAGATCTTGGACAGAGAAAAG 116480

Oy 1912 ATCAGTATGAAATCTATGACGTCTGATATAGAAATCTGAGATATTTTAAACAGTACTG 1971
 |||||
 Db 116479 ATCAGTATGAAATCTATGACGTCTGATATAGAAATCTGAGATATTTTAAACAGTACTG 116420
 Oy 1972 TTGATTTCTTAATCTTGAACAAATATGACAGGGTAAATGATGATTAACGTAGAGAAC 2031
 |||||
 Db 116419 TTGATTTCTTAATCTTGAACAAATATGACAGGGTAAATGATGATTAACGTAGAGAAC 116360
 Oy 2032 TGAACCTGGTGAAGGCTATCTGATGAAATCTCTGATATCTTACCAATTTTGGGTAG 2091
 |||||
 Db 116359 TGAACCTGGTGAAGGCTATCTGATGAAATCTCTGATATCTTACCAATTTTGGGTAG 116300
 Oy 2092 TCTAGAAAGCAATGCAAAATTAAGTATCTTGAAGAAAAA 2141
 |||||
 Db 116299 TCTAGAAAGCAATGCAAAATTAAGTATCTTGAAGAAAAA 116250
 RESULT 7
 HUMTPX1A
 LOCUS HUMTPX1A 1335 bp mRNA linear PRI 14-JAN-1995
 DEFINITION Human testis-specific protein (Tpx-1) mRNA, complete cds.
 ACCESSION M25532 J04741
 VERSION M25532.1 GI:339882
 KEYWORDS testis-specific protein.
 SOURCE Human adult testis, cDNA to mRNA, clones H4-1 and p3-1.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1335)
 Kasahara, M., Gutknecht, J., Brew, K., Spurr, N. and Goodfellow, P. N.
 Cloning and mapping of a testis-specific gene with sequence
 similarity to a sperm-coating glycoprotein gene
 Genomics 5 (3), 527-534 (1989)
 JOURNAL MEDLINE 90129048
 PUBMED 2613236
 COMMENT Draft entry and computer-readable sequence for [1] kindly submitted
 by M. Kasahara, 09-JUN-1989.
 FEATURES
 source location/Qualifiers
 1..1335
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="6p21-qter"
 1..1335
 /gene="GAPDL5"
 <1..1335
 /gene="GAPDL5"
 /product="testis-specific protein mRNA"
 190..921
 /gene="GAPDL5"
 /note="testis-specific protein precursor"
 /codon_start=1
 /protein_id="AA61220.1"
 /db_xref="GI:339883"
 /db_xref="GDB:G00-119-257"
 /translation="MALLPVLEFVLTLLPSLPARKDPAFLITLOLOVREIVNKH
 NELKRAVSPPRSNNLKKHEKSEVTTMORNRANKCTLDHSDPEDRKSTRCGENLIMS
 DPISWSSAIOSWDEILDFYVGVGPKSPNAVGHYTLQVWSTYGVCGGLAYCPNDOS
 LKYYVQOYCPAGNNMNRNKPYYOQGPACGPDGDKLCTNCSQYODLSDNSDLK
 NTACSEHLELKEKCAICLENKIY"
 190..249
 /gene="GAPDL5"
 /note="testis-specific signal peptide (put.): putative"
 mat_peptide 250..918
 /gene="GAPDL5"
 /product="testis-specific protein (put.): putative"
 BASE COUNT 428 a 279 c 271 g 355 t 2 others
 ORIGIN Chromosome 6.
 Query Match 24.0%; Score 515.2; DB 9; Length 1335;
 Best Local Similarity 68.8%; Pred. No. 4,4e-85;
 Matches 798; Conservative 0; Mismatches 323; Indels 39; Gaps 5;
 Oy 1 TGATGAACAATACTTACTCTGCTGGAACCA---CTGCATGACATTTATTCACA 56

D	171	TAATAAGATGATATTTTCATCTGCTCAGAAAAACCAATTTCCAGCAATGGCTTTACTA	230
Q	57	GTGCTGTTGTTCCGGTTCGTGGGCTGCTTCCATCTTTTCCAGCAATGATATAGAT	116
D	231	CCGGTGTGTTTCTGGTTACTGCTGCTCCATCTTTACCTGCA---GAAAGAAAGAT	287
Q	117	CCGCCTTTACTGCTGTTGTTAACCCACCACCAACAGTAGTCCAAAGGAGATGTGATAG	176
D	288	CCCGCTTTTACTGCTTTGTTAACCCACCAGTTCGAAAGTGCAAAGGAGATTTAATATA	347
Q	177	CACAATGAGTGAGAGAGCAGTATCTCCCCCTGCAGAAACATGCTGAGATGGAAATG	236
D	348	CACAATGAACTAAGGAAGCAAGTCTCCACCTGCCAGTAAACATGCTAAAGATGGAATG	407
Q	237	AACAAAGAGCTGCAGCAAAATGCCCAAAGTGGGCAACCAAGTCGCAATTACAGACAGT	296
D	408	AGCAGAGAGGTAAACAAAGATGCCCAAAGTGGGCAACCAAGTCGCTTACAAATATGT	467
Q	297	AACCAAGAGTGAATGACAAAGTCTAAATGGGAGAACTCTACTGTCAGTAGTCC	356
D	468	GATCCAGAGACCCGCAAAACACAGTACAGATGGTGGAATACTCTAATGTCAAAGTAC	527
Q	357	CCCACTCATGCTCCACAAAGTCGCTTGGATGATGAGTACATATTTGACTTT	416
D	528	CCCTACTCTGGTCTTCTGCAATCCAAAGTGGTATAGAGAGATCTTAATTTTGTCTAT	587
Q	417	GGGTAGGGCCAAAGACTCCCAACGAGATGGTGGACATTTACACAGGTGTTGGTAC	476
D	588	GGGTAGGACCAAGAGTCCCAATGCAATGTTGGACATTTATACAGCTTGTGGTAC	647
Q	477	TCTTCATACCTCGTGGATGGAATGGCTACTGTCCCAATCCAAAAGTTCTAAATAC	536
D	648	TGCACTTACAGGAGAGCTGTGAATGGCTACTGTGCCAATCAAGATGTCTAAATAC	707
Q	537	TACATTTTGGCCAAATTTCTCGTGGTAAATGGGCTAAATAGACTATATGCTCTAT	596
D	708	TACTATTTTGGCCAAATTTCTCGTGGTAAATATATATATGAAAGATACCCCGTAC	767
Q	597	GAAACAGAGACCTTGTGCCAGTTGCCAGATACGTGAGCATGAGTATGACACAT	656
D	768	CACAAAGAAACACCTTGTGCCGAGTGCCTGATGACTGTGACAAAGGACTATGACACAT	827
Q	657	GGTTCGAATGACGAAGTCTCTATAGTATGTAATGTAAGTTGAAGTCACTTAACCTGT	716
D	828	AGTTGCAGATCAAGATCTCTTAAGTATACGTGATTCCTTAAGATATCACCTGGCTGT	887
Q	717	AACATCAGTTGGTCAGAGCAGTGTCCAGAGCCTCTGCAATTTGTTCAACAGCATTTA-	775
D	888	GAACTATAGTTACTCTCAGAGAAAGTGCAAAGGCTACTTGGCTTATGAGAACCAAAATTTAC	947
Q	776	TTAAATACGATTAACACACCAGTAGGCGTATGTAGAGAGGAGTAGATATATCTACTAG	835
D	948	TGATTTACCTAGTAGAGCATGTTGCAGAGCTGATGATGAAGGCTGACATCTTAAT--	1005
Q	836	ATTGGCATCTACTAGATTTAACTATATACCTGAGAAATTTGATGGCATGTTGATAC	895
D	1006	-----GCGACATACCACTGAGAAATTTGATGATGTTGTGTGAC	1042
Q	896	ACATTTGATTTCAAAATGTTTCTCTGAGTCTGCTTTTATTTTACAAAATATTTTTC	955
D	1043	AAATTTATTTTCAAAAGAGCATATCTTCTCCCCAGATCATACAGAGAAACACTTTTCA	1102
Q	956	ATACAAATGTTAAAGAAAGAAACAAATCTATACACAACTTTGGATTTTATATATAA	1015
D	1103	GGCAATGATTTACAAAGATGAGCATAGTAGATGATGACAACTGTGAACCTGTGACATAA--	1166
Q	1016	CTTTGCTATTTAAATTTACTGCAATTTAATTTAGGGTGAATAATTTGAAGTTGATTTCTCA	1075
D	1161	-----TTTAGCTTTAATACGAATCAATCAGTTGTGAGATTTTGAACCTGTATACCA	1216
Q	1076	TATGACTAAGTCACTAAACCTGTGATGGAAGTGAATAATATGTTCTCTACAAACAAT	1139
D	1217	TAGATTTAGTCACTAGACCTTTGGATCAAAATGGTGCATTAAGATTTCTGTAAACAT	1276

RESULT 9											
LOCUS	LOCUS	1406 bp	RNA	linear	PRI 12-APR-1996						
DEFINITION	H.sapiens mRNA for cysteine-rich secretory protein-2/type I.										
ACCESSION	X95239										
VERSION	X95239.1	GI:1262816									
KEYWORDS	CRISP-2 gene; cysteine-rich secretory protein-2/type I.										
SOURCE	Homo sapiens.										
ORGANISM	Homo sapiens										
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.										
AUTHORS	1 (bases 1 to 1406)										
TITLE	Kratzschmar, J., Haendler, B., Eberspaecher, U., Roosterman, D., Donner, P. and Schleunig, W.D.										
JOURNAL	The human cysteine-rich secretory protein (CRISP) family. Primary structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3										
MEDLINE	Eur. J. Biochem. 236 (3), 827-836 (1996)										
PUBMED	8665901										
REFERENCE	2 (bases 1 to 1406)										
AUTHORS	Haendler, B.										
TITLE	Direct Submission										
JOURNAL	Submitted (18-JAN-1996) B. Haendler, Schering AG, ICMB, S109/517, 13342 Berlin, FRG										
COMMENT	Related sequence J04741.										
FEATURES	Location/Qualifiers										
source	1. 1406										
	/organism="Homo sapiens"										
	/db_xref="taxon:9606"										
	/clone_id="testis cDNA library"										
gene	241..972										
	/gene="CRISP-2"										
CDS	241..972										
	/gene="CRISP-2"										
	/codon_start=1										
	/product="cysteine-rich secretory protein-2/type I"										
	/protein_id="CAA64526.1"										
	/db_xref="GI:1262817"										
	/db_xref="SWISS-PROT:P16562"										
	/translation="MALPPLPFLVTVLLPSLPAGKRPATALLTOLQVORELYNKH NELKRAVSPASNMLKMEWSREVTNQRANKCTLOHSDPEKDKISTRGENDLYKSS DPDSSMSAIQSWYDEILDFTVGVGPKSPNAYVGHYIOLVWSTYVCGCIAYCPNDLS LKYYVCOAYCPAGNNMNRKNTPYOOGCPACGCPDDDKGLCTNSCOYODLSMCDLSLK NTACGHEHLEKKEKATCLCENKRY"										
sig_peptide	241..300										
	/gene="CRISP-2"										
polyA_signal	1362..1367										
BASE COUNT	458 a 293 c 289 g 366 t										
ORIGIN											
Query Match	24.0%	Score 513.6	DB:9	Length 1406							
Best Local Similarity	68.7%	Pred. No. 8.5e-85									
Matches 797	Conservative 0	Mismatches 344	Indels 39	Gaps 5							
1	TGATGAAACAATATCTTCATCCTGCTGGAACCA---CTGCAATGCATTATTCCCA	56									
Db	193 TAATAAGTAGTATTTCATCCTGCTGGAACCAACCAATTTCCAGCAATGGCTTACTA	252									
0y	57 GTGCTGTTGTCCTGCTGCTGGGCTGCTTCATCTTTCCAGCAATGAATAGAT	116									
Db	253 CCGGCTGTTGCTGCTACTGTGCTGCTTCATCTTACTTCGCA---GAAGGAAGAT	309									
0y	117 CCGGCTTACTGCTTTTAACCAACCAACAAGTGAAGGAGAGATGTGAATAG	176									
Db	310 CCGGCTTACTGCTTTTAACCAACCAAGTGAAGTGAAGGAGATGTGAATAA	369									
177	CACAATGAAGTGAAGAGAGATATTCGCCCTGCGCAACCAATGCTGAAGATGAATGG	236									

```

Db 370 CACATGATAGTAAAGAAAGCACTCTCTCCACCTGCCAGTAAACATGTTAAAGATGGAATGG 429
Oy 237 AACAAAGAGGCTGCAGCAAAATGCCCAAAAGGGGCAAAACCAAGTCAATTCAGACACAGT 296
Db 430 AGCAGAGAGGTAAACAAGAAATGCCCAAAAGGGGCAAAACCAAGTCAATTCAGACACAGT 489
Oy 297 AACCCAAAGATGATGACAAAGTCTAAATGTGTGAGAAATCTCTACATGTGCAGTGGC 356
Db 490 GATCCAGAGGACCGCAAAACCAAGTCAATGTGTGAGAAATCTCTATGTGTCAAGTGC 549
Oy 357 CCCAGCTCATGTGTCAACCAAGTCAATGTGTGTGATGATGATGATGATGATGATGATGAT 416
Db 550 CCTACTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 609
Oy 417 GGTGAGGGCCAAAGACTCCCAAGCAGTGTGTGACATTAATACAGAGTGTGTGTGTAC 476
Db 610 GGTGAGGACCAAAAGAGTCCCAATGCTGTGTGACATTAATACAGAGTGTGTGTGTAC 669
Oy 477 TCTCATACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 536
Db 670 TCGACTTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 729
Oy 537 TACTATGTTGCCAATATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 596
Db 730 TACTATGTTGCCAATATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 789
Oy 597 GAACAAGAGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 656
Db 790 CAACAAGAGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 849
Oy 657 GGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 716
Db 850 ACTTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 909
Oy 717 AACATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 775
Db 910 GAACATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 969
Oy 776 TTAATACCATTTACACACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 835
Db 970 TGAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1027
Oy 836 ATTGGCATCTAATGATTTAAATATATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 895
Db 1028 -----GCCACATPACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1064
Oy 896 ACATTTGATTTCAAAATGTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 955
Db 1065 AAATTTGATTTCAAAAGAGCAATGCAATCTTCTCCCCAGATTCATCAGAAATCACTTTGA 1124
Oy 956 ATACAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1015
Db 1125 GGCATGATTTCAAAAGAGCAATGCAATCTTCTCCCCAGATTCATCAGAAATCACTTTGA 1182
Oy 1016 CTTTGTGATTTAAATTTTACTGATTTAAATTTAGGAGTAAATTTGAAAGTGTGTGTGTGTGT 1075
Db 1183 -----TTTACTGCTTTTAAACGAATCAATCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1238
Oy 1076 TATGACTAAGTTCATTAACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1135
Db 1239 TAGGATTTGCTACTAGAGCTTTGGATCAAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1298
Oy 1136 GTACAAAAAGAACATATA 1155
Db 1299 GCTAAGAAGAGAGCTGTA 1318

```

```

RESULT 10
ECRISP3 ECRI3P3 1295 bp mRNA linear MAM 03-FEB-1998
LOCUS DEFINITION Equus caballus mRNA for cysteine-rich secretory protein-3.
ACCESSION AF001400

```

```

VERSION AJ001400.1 GI:2388784
KEYWORDS CRISP-3; cysteine-rich secretory protein-3.
SOURCE horse.
ORGANISM Equus caballus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE 1 (bases 1 to 1295)
AUTHORS Schambony, A.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1997) Schambony A., Institut fuer
Reproduktionsmedizin, Tierärztliche Hochschule Hannover, Bunteweg
15, D-30559 Hannover, GERMANY
REFERENCE 2 (bases 1 to 1295)
AUTHORS Magdalen, L., Gassel, M., Vree, J., Schambony, A.M., Urbanke, C.,
Raida, M., Topfer-Petersen, E. and Calvete, J.J.
TITLE Biochemical and conformational characterisation of HSP-3, a
stallion seminal plasma protein of the cysteine-rich secretory
protein (CRISP) family
JOURNAL FEBS Lett. 420 (2-3), 179-185 (1997)
MEDLINE 98119394
PUBMED 9459306
FEATURES
source Location/Qualifiers
1..1295
/organism="Equus caballus"
/db_xref="taxon:9796"
/sex="Male"
/life_stage="adult"
/dev_stage="adult"
/gene="HSP-3"
/cds="crisp-3"
59..796
/gene="crisp-3"
/codon_start=1
/product="cysteine-rich secretory protein-3"
/protein_id="CAA04729.1"
/db_xref="GI:2388785"
/db_xref="E:2388785"
/translation="MALPVLFLVLAIVLPPASGDPGFAISITSEVOKETVNH
HNDLRRTVSPLASMLKQWDBSKTITNQNANKLLDHSKEDRAVDTMKGELPM
SSIPMSDAIDNWHDEVHDFRYGVGPVPAVGHVYVWVYSYRVGCCIAICPKO
GLKRYVCOYCPAGNVNKNITPVEOGTPCARCGNCDNGICNSCEBLSNCD
LKRIGCEHELLENCKRTKCCENKRY"
59..124
sig_peptide
polyA_signal /gene="crisp-3"
BASE COUNT 428 a 257 c 245 g 365 t
ORIGIN
Query Match 23.6%; Score 505.8; DB 4; Length 1295;
Best Local Similarity 70.5%; Pred. No. 2.4e-83;
Matches 803; Conservative 0; Mismatches 297; Indels 39; Gaps 8;
Oy 4 TGAACAAATACTTCTCTCTGTGTGAAACCACTGCAATGACATTAATCCAGTCTGT 63
Db 24 TAAAGAGATGCTTCACTGTGTGTGAGAACCA---ACATGGCATTAATACAGTGTGG 80
Oy 64 TGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 123
Db 81 TGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 137
Oy 124 TTAAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 183
Db 138 TTGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 197
Oy 184 AACTGAGAGAGAGATGTCTCCCTGTGCAAGAAACATGTGTGAGATGATGAGAACAAAG 243
Db 198 ACCTAAGAGAGAGAGATGTCTCCCTGTGCAAGAAACATGTGTGAGATGATGAGAACAAAG 257
Oy 244 AGGCTGACGAATATGCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 303
Db 258 AGACAGCAACCAATATGCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 317
Oy 304 AGATCG---AATGACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360

```

```

Db 318 AAGATCGCGCAGTCGTCACATGAAATGTGTGAGAAATCTTTTATGTCGACTAATACCTA 377
Oy 361 GCTCATGCTCAACAACCAATCCAAAGCTGTTGATGATGATGATGATGATGATGATGATG 420
Db 378 ATTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 437
Oy 421 TAGGCGCAAGAGCTCCCAACGAGTGTGGACATTAACAGAGTGTGTGTGTGTGTGTGTGT 480
Db 438 TGGCGCCCAAGAGCTCCCAATGAGTGTGGACATTAACAGAGTGTGTGTGTGTGTGTGTGT 497
Oy 481 CATACCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 498 TTTACCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 557
Oy 541 ATGTTTGCATATATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 558 ATGTTTGCATATATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617
Oy 601 AAGAGCACTTTGTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 618 AAGAGCACTTTGTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
Oy 661 GCAATAGAGAGATCTCTATAGTACTGTAAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 720
Db 678 GCGATAGAGAGATCTCTATAGTACTGTAAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 737
Oy 721 ATCATGTTGTCAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 780
Db 738 ATGATGTTGTCAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 797
Oy 781 TACGATTACACACCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 840
Db 798 GTTCAGTGTGATGTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 834
Oy 841 GCATCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Db 835 GCTGATCTCTGTTGTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 893
Oy 901 TGATTTCAATATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 894 TGATTTCAATATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
Oy 961 AATGTTTCAATATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 952 ATGATTTCAATATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008
Oy 1021 TGATTTTCAATATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1009 TACTTTTCAATATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1067
Oy 1081 CTAACTTCAATATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139
Db 1068 CTAACTTCAATATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126

```

```

MEDLINE 96354287
PUBMED 9115720
REFERENCE 2 (bases 1 to 1388)
AUTHORS Foster,J.A. and Gerton,G.L.
TITLE Direct Submission
JOURNAL Submitted (08-SEP-1995) James A. Foster, Department of Obstetrics
and Gynecology, Division of Reproductive Biology, University of
Pennsylvania, 36th and Hamilton Walk, Philadelphia, PA 19104, USA
FEATURES
source
1. 1388
/organism="Cavia porcellus"
/strain="Hartley"
/db_xref="taxon:10141"
/cell_type="sperm acrosome"
/tissue_id="testis"
185..919
/codon_start=1
/product="autoantigen 1"
/protein_id="AAC52616.1"
/db_xref="GI:1016756"
/translation="MALPVPVFLITMLPCVLTKGKDPFAFTALITQSQVONEINK
HNOLRKSVTPSPASMLKMEKNSREAAVNAOKANRCTLVHSNPDRLKSTGKENVLS
SDPSMSDAIOSWPEDESODPEFGVSHNAVYGHYOLVWYSYLWGCGGLAXPNOD
SLKTYVCOICPACNNVYTKNTPYKQGIPCASCPGHCENGLCTNSCEYEDLNSCESL
KNTAGCEHQLVLEKCKATCRCEKLY
247..309
/note="encodes region previously determined by
microsequencing (Hardy et al. Biol Reprod Vol. 38, p.
423-437, 1988), PIR Accession Number A60802"
BASE COUNT 444 a 276 c 285 g 383 t
ORIGIN
Query Match 22.2%; Score 476.2; DB 10; Length 1388;
Best Local Similarity 68.8%; Pred. No. 6.5e-78;
Matches 758; Conservative 0; Mismatches 303; Indels 40; Gaps 6;
Oy 1 TGATGAACAATATCTGATCTGCTGGAACCA-----CTGCATGACATTATTC 53
Db 137 TAATTAACAGATATGATCTGCTGTGAGAACCAATTCCTCCAGCAATGCGCTTACTC 196
Oy 54 CCAGTGTGTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 113
Db 197 CCAATGTTGTTCTTCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 253
Oy 114 GATCCCGCTTTTACTGCTTTGTTAACCAACCAACCAAGTGCAGATGTTGTAAT 173
Db 254 GATCCAGATTTTACCGCTTTGTTAACCAACCAACCAAGTGCAGATGTTGTAAT 313
Oy 174 AAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 233
Db 314 AAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 373
Oy 234 TGAACAAGAGGCTGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 293
Db 374 TGAACAAGAGGCTGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433
Oy 294 AGTAACCAAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 353
Db 434 AGTAACCAAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 493
Oy 434 GATTAATCCAGAGATGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 493
Oy 354 GATTAATCCAGAGATGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 413
Db 494 GATTAATCCAGAGATGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 553
Oy 414 TTTGTTAGGAGGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 473
Db 554 TTTGTTAGGAGGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 613
Oy 474 TACTTCTCATCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 533
Db 614 TATTCGTTATGTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 673
Oy 534 TACTTCTCATCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 593

```

RESULT 11
 CPU35712 1388 bp mRNA linear ROD 18-JUN-1996
 LOCUS Cavia porcellus acrosomal autoantigen 1 mRNA, complete cds.
 DEFINITION U35712
 ACCESSION U35712
 VERSION U35712.1 GI:1016755
 KEYWORDS
 ORGANISM Cavia porcellus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Hystriognathli; Cavidae; Cavia.
 (bases 1 to 1388)
 REFERENCE Foster,J.A. and Gerton,G.L.
 AUTORS Autoantigen 1 of the guinea pig sperm acrosome is the homologue of
 mouse Tpx-1 and human TPX1 and is a member of the cysteine-rich
 secretory protein (CRISP) family
 TITLE Mol. Reprod. Dev. 44 (2), 221-229 (1996)

JOURNAL


```

Db 674 TACACTACGTTTCCATACCTGCTGCTGTATATATGTATACAAAGAACACTCTCT 733
OY 594 TATACACAGAGACACTTGTGCCAGTTGCCAGATTAAGTACAGATGACTATGACC 653
Db 734 TATAAACAGAAATACCTGTGCCAGCTGCCCTGACATTTGTGAAAATGAGACTATGACC 793
OY 654 AATGGTTGCAAGTACGAAGATCTATATAGTACGTAAAGTTGAAGCTCACAATTAC 713
Db 794 AACAGTTGTAGTAAAGATCTCTTAGTAACTGCGATCTTGAACAAATACAGCTGC 853
OY 714 TGTAAACATCAGTTGTGTCAGGACAGTTCGAAGCATCTGCAATTTGTTCAACAGCAT 773
Db 854 TGTGAACATCAATTTGCTGTGAAAAGTCAAGGCTACTTGCAGATGTGAAGATTAATA 913
OY 774 TATTAAATAGCATTAACACACCCAGTAGGCTATGTAGAGAGAGACTATATCTACTT 833
Db 914 TACTAAA-----ACTCACCCGCTTTGTGCAGAGACTA-----GTGG 946
OY 834 AGATTGGCATCTACTAGTATTAACATATATAGCTGAGAAATGTAGGATGTTGAT 893
Db 947 AAACAGGTGACACCTTTGTTGTGACAAACAGATTAACCTCTGATAGGTTAGAT 1006
OY 894 ACACATTTGATTTCAATGTTTCTCTGATCTGCTTTTATTTTACAAAATATTTT 953
Db 1007 ATAAGTTAGTATGCAAAATAGAGGTACTACATAGTCTCTGATATTTGAAAAATTTGAT 1066
OY 954 TCATACAAATGTTTAAAGAAACAAATCTATACAAACACTTTGATTTTATATATA 1013
Db 1067 TTGTACACCAATTTTCAAAAAGAACATTTGCT-ATAACATCTGTGACCTTTGATAGAGC 1125
OY 1014 AACTTGTGATTTAAATTTAGCAATTTAATAGGTTGAAAATTTGAAAGTTGATTTCT 1073
Db 1126 A--TTCAATCTTTAAATTTAATGATTTGATATCAATCAGATTTTGAAGGCTGACACT 1183
OY 1074 CATATGACTAAGTTCATAAA 1094
Db 1184 CATAGGACTTAGGTCACCTAGA 1204

RESULT 12
MUSTPX1A 1418 bp mRNA linear ROD 27-APR-1993
LOCUS MUSTPX1A
DEFINITION Mouse testis-specific protein (Tpx-1) mRNA, complete cds.
ACCESSION M25533.1 J04741
VERSION M25533.1 GI:202126
KEYWORDS testis-specific protein.
SOURCE Mouse (strain BALB/c) testis, cDNA to mRNA, clone B4(1)4.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1418)
AUTHORS Kaashara, M., Gutknecht, J., Brew, R., Spurr, N. and Goodfellow, P. N.
TITLE Cloning and mapping of a testis-specific gene with sequence
similarity to a sperm-coating glycoprotein gene
JOURNAL Genomics 5 (3), 527-534 (1989)
MEDLINE 90129048
PubMed 2613236

COMMENT
Draft entry and computer-readable sequence for [1] kindly submitted
by M. Kaashara, 09-JUN-1989.
Location/Qualifiers
FEATURES
SOURCE
1..1418
/organism="Mus musculus"
/db_xref="taxon:10090"
<1..1418
/product="testis-specific protein mRNA"
214..945
/note="testis-specific protein precursor"
/codon_start=1
/protein_id="AAA0472.1"
/db_xref="gi:202127"
/translation="MAFQVMEFVALLRSPLEGRKDPFTSLTNOLOYOREIVNK
HNELRSVNPITGSDILKMEWSIQATITNQNQKANCILIEHSSDKDKINIRGENLYMS

```

QY 936 ATTTACAAAATATTTTCATACAAATGGTTAAAAAGAACAAATCTATACAAAC 995
 Db 1076 TTCAATCATATATATCTGACAAACGACCATTTAAAGGTTGCTATAGTATACAAAC 1135
 QY 996 TTTCATTTTATATATACTGTTGATTTAATTTACTGATTTAATTTAGGCGAA 1055
 Db 1136 TTTCGATTTGCA---CCAACTTGTCTGCTTAAATATAGTAAAGTAAAGTGA 1192
 QY 1056 TTTTGAAGTTGTATTTCTATATGACTAAGTCACTAAACCCCTGGATGAAAGTGA 1115
 Db 1193 TTTTGAAGTTGTACACATATAGCTCATCACTGACCTGGAACCTTTAAAAAGAA 1252
 QY 1116 TTATGTTCTGACAAATATGACAA 1141
 Db 1253 TATTTGTCTGCTAAATTAACCTGCA 1278
 RESULT 13
 AB009662
 LOCUS AB009662 1280 bp mRNA linear ROD 01-AUG-1998
 DEFINITION Rattus norvegicus mRNA for testis specific protein, complete cds.
 ACCESSION AB009662
 VERSION AB009662.1 GI:3374579
 KEYWORDS testis specific protein.
 SOURCE Rattus norvegicus (strain:Donryu) 20 days after birth male Testis
 Spermatogenic cells cDNA to mRNA, clone:TpX-1.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 1280)
 Meda, T., Sakashita, M., Ohba, Y. and Nakanishi, Y.
 Molecular cloning of the rat TpX-1 responsible for the interaction
 between spermatogenic and Sertoli cells
 Biochem. Biophys. Res. Commun. 248 (1), 140-146 (1998)
 2 (bases 1 to 1280)
 Nakanishi, Y.
 Direct Submission
 Submitted (11-DEC-1997) Yoshiobu Nakanishi, Kanazawa University,
 Faculty of Pharmaceutical Sciences; 13-1 Takara-machi, Kanazawa,
 Ishikawa 920, Japan (Tel:076-234-4480, Fax:076-234-4480)
 FEATURES
 source
 1. 1280
 /organism="Rattus norvegicus"
 /strain="Donryu"
 /db_xref="taxon:10116"
 /clone="TpX-1"
 /sex="male"
 /cell_type="Spermatogenic cells"
 /tissue_type="Testis"
 /dev_stage="20 days after birth"
 78. 809
 /codon_start=1
 /product="testis specific protein"
 /protein_id="BA012029.1"
 /db_xref="GI:3374580"
 /translation="MAFVQVLFVFAVLLPPTGKDPFATLTTNOIOVREIILK
 HNELRQVSPGNSILKMEVNOAANAQMANCIHESSTEDRKINICGSELYLS
 TDPTNRTVIOSENEFVFGVAPNVAHYTLQVWSSPKVCGVAYCPNDOT
 LKRYFYCHVCPGNVWKSTPYHOGPCASCRNNDNGICTNSCPREDLLSNCEISK
 SASGCKHELKACENTCLEDKIH"
 1280
 polyA_site
 /note="20 a nucleotides"
 BASE COUNT 416 a 263 c 259 g 342 t
 ORIGIN
 Query Match 17.28; Score 367.8; DB 10; Length 1280;
 Best Local Similarity 61.88; Pred. No. 5.9e-58;
 Matches 681; Conservative 0; Mismatches 387; Indels 37; Gaps 5;
 QY 37 CTGCATGACATATATCCAGCTGTTGTTCTGTGCTGGGCTGCTTCATCTTTTC 96

Db 73 CAGCCATGGCTGGTTCAGAGTATGCTGTGTTCTGCTGCTGACCA---TTCC 129
 QY 97 CAGCAATGAGATATAGATATCCCGCTTTTACTGCTTTGTATACCAACCAACAGTGC 156
 Db 130 CACCACAGAGAGAAAGGATCCAGACTTGTCTTTGACAAACCAAAATACAAAGTTC 189
 QY 157 AAAGGAGATTTGTAATAGACAAATGAACTAGAGAGAGACATATCTCCCTGCGAGAA 216
 Db 190 AAAGAGAGATCTACTTAACAAATGAACTAGAGAGAGAGAGATGAGCCCTGCGACAA 249
 QY 217 ACATCTGAGATGGAATGGAACAAAGAGCTGACGCAATGCCAAAAGTGGCAAC 276
 Db 250 ACATCTGAAATGGAATGGAACGATACAGAGCAGCAAAATGCTCAAAAGTGGCTAATA 309
 QY 277 AGTGAATTAAGACACAGTAAACCAAGAGATGCAATGCAAGTAAATGCTGTAGA 336
 Db 310 ACTGATTTTATGACACAGATATACAGAGACCGGAAATCAATATTAATGTGGCGAGA 369
 QY 337 ATCTCTACATGTCAGATGCCCCCAGCTCATGCTCACAGCAATCCAAAGTGTGTGATG 396
 Db 370 ATCTCTATATGTCGACTGACCTCATCTCTGAGAGAACGTAATTAAGTGTGTAGAG 429
 QY 397 AGTACATGATTTTACTCTTGTGTAGGCGCAAGACTCCCAAGCGACTGTGACAT 456
 Db 430 AAAATGAAACACTGCTTTTGGCGCTAGAGACTTA---ACCCAAATCCGCTGCGAGACT 486
 QY 457 ATACACAGTGTGTGATGCTTCTCATACCTGCTGATGATGGAATGCTGACTGTCCCA 516
 Db 487 ACATCTGACTGTTGTGATATTCATTTTCAAGTTGGATGTGAGATCTTCTACTGTCCA 546
 QY 517 ATCAAAAAGTCTAAATACTATATGTTGGCAATATGTCCTGCTGTAATGGGCTA 576
 Db 547 ATCAAGATACCCCTGAAATACTCTATGTTGGCATTTACTGTGCTTAACCAAGCTGA 606
 QY 577 ATAGCTATATGTCCTTATGAAAGAGACACCTGTGCGCAAGTGTCCCAATATCTGT 636
 Db 607 TGAAGAGATGACCCATATCTCAAGAGACACCTGTGCTAGTGTGCCAATTAACCTGT 666
 QY 637 AGATGACTATGACCAATGATGTCAGATGAGATGATGATGATGATGATGATGATG 696
 Db 667 ATATGATGATGTCGACCAATAGCTGTGATTTGGAAGATTTACTAGTACGTGAATCT 726
 QY 697 TGAAGCTTACATTAACCTGTAAACATCATGTTGTCAGGAGACAGTGTGACAGCAT 756
 Db 727 TGAAGAGTTCAGAGGCTGTAAACATGATGCTCAAGCAAGTGTGAGGCTACTGCC 786
 QY 757 ATGTCACAGCATTTATTAATATAGCATTTACACACCGATGAGCTATGTAGAGAG 816
 Db 787 TATGTGAAGCAAAATTTATTAACATGCCAGCGTGCAGCATGACATCATAGAGAG 846
 QY 817 AGTCAGATTTATCTACTTATGATTTGGCATCTACTTAGATTTAACAATATCTAGTGA 876
 Db 847 GGGTACAGACTTAGTTGACATAGCA-----GGGAAA 880
 QY 877 TTGTAGGATTTTGTATACATTTGATTTCAATGTTTCTCTGATGCTGTTT 936
 Db 881 CCTATAGAGAGTATGTAACAGTGCATCCCAATATGACAAAGCTTCTCTCTGAT 940
 QY 937 TTTTACAAAATATTTTTCATACAAATGCTTAAAGAAAGAAACAAATCTATACAAAC 996
 Db 941 TTATATGAAATGCTTTTCAATACAGCATTAAGAAAG--GTGATATTAAGATTAAC 998
 QY 997 TTGCATTTTATATATACTTGTGATTTAATTTACTGATTTAATTTAGGCGAAAT 1056
 Db 999 CTGATTTTGA---CCAACTTGTCTGCTTCAAAATGATGAGGAATCAAGTGTGAGAT 1055
 QY 1057 TTTGAAAGTTGATTTCTCATATGACTAAGTCACTAAACCCCTGATGAAAGTGA 1116
 Db 1056 TTTGAAAGTTGATCAATATGCTGATCATCTGATGAAAGTGAAGAGAGAGACTGT 1115
 QY 1117 TATGTTCTGACAAATATGACAAAGAA 1147

AUTHORS Charost, N.J., Joseph, D.R., Wilson, E.M. and French, F.S.
 TITLE Molecular cloning of complementary deoxyribonucleic acid for an androgen-regulated epididymal protein: sequence homology with metalloproteins
 JOURNAL Mol. Endocrinol. 2 (10), 999-1004 (1988)
 MEDLINE 89039913
 PUBMED 2460753

FEATURES
 source location/Qualifiers
 1. 1498
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue-type="epididymus"
 /dev_stage="adult"
 1. 1498
 /gene="AEG"
 <1. 1498
 /gene="AEG"
 /product="alternative mRNA"
 join(<1. 918,1457. 1493)
 /gene="AEG"
 /product="alternative mRNA"
 11. 851
 /gene="AEG"
 /codon_start=1
 /product="epididymal glycoprotein"
 /protein_id="AAB59716.1"
 /db_xref="GI:202773"
 /translation="MALMLYLFLAALVPSLLODTTDEMDELNTSTKLVSQEEI
 INKNQRLRTVSPSGDLRYEMDHDAVYNAOKYKAVHYHTQVYNSPLVACGAEEDP
 FMANYPASMSVSDYDDESLDFVFGPKYGVKAVHYHTQVYNSPLVACGAEEDP
 DQPKYFVCHVCGNVRGLVSPYRGEPCDPCNCEDEDTNSG
 DLKMSCDPLKEGCRASCFEDKH"
 1476. 1477
 /gene="AEG"
 /note="at in long clone; aat in short clone"
 /replace="aat"
 1490
 /gene="AEG"
 /note="t in long clone; c in short clone"
 /replace="c"
 variation
 BASE COUNT 464 a 289 c 271 g 474 t
 ORIGIN

Query Match 16.98; Score 362.2; DB 10; Length 1498;
 Best Local Similarity 62.38; Pred. No. 6,1e-57;
 Matches 713; Conservative 0; Mismatches 403; Indels 29; Gaps 8;

```

OY 11 AATCTTCATCCTGCTCGAAGAACACATGACATGATTTCCACAGCTGTTGTTCT 70
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 AATATCTCATTTCTGCTCT-GAAATAATAGAACATGCGATTAATGTTAGCTGTTGTTCT 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 71 GGTTCGCTGGCTGCTTCATCTTTCCAGCAAA-----TGAAGTAAGATCCGCTTT 124
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 140 GGTGCTGCTATGCGACCATCTCTCTTCAGATACCACTATGATGGATAGAGACT 199
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 125 TACTGCTTTGTTAAACCAACCAACAGTGAAGAGATTGTAATTAACCAATGA 184
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 200 TGAGAAATTTGCAACCACTAAAGTGTCAAGAAAGATCATTAACAACCAACCA 259
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 185 ACTGAGAGACAGATATCTCCCTCCAGAAACATGCTGAAGATGGAATGGAAACAAGA 244
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 260 ATTAGACGAAGAGTTTCTCGTCTGTAGTACTTACTAAGAGTGGAGACATGA 319
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 245 GGTGCAAGCAATGCCAAAAGTGGCAAAACAGTCAATTAACAGACATAACCCAAA 304
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 320 TGTATTATGTGAAGCTGAGAAATGGGCAACAGGTGCAATTACATCATCTCTTACA 379
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 305 GGATCGAATGCAAGTCAAAATGTGTGAGAAATCTTACATGTCAAGTGCCTCCAGCTC 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 380 ACAGAGACCAACCATTAATAATGTGTGAGAAATTTGTTACATGCAATTAACCTGCATC 439
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 365 ATGTCACAAGCAATCAACACTGTTTGATGATGATACATGATTTTGACTTTGGTGTAG 424

```

```

Db 440 GTGGTCTCTCTGTAAATCCAGATGGTATGATGAATCCCTTATTTGCTTGGTTGG 499
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 425 GCCAAGACTCCCAACGAGAGTGTGGACATTAATACAGAGTGTGGTACTCTTCATA 484
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 CCCAAAAAAGTTGGTGTAAAGTGGACACTATATCTAGGTGTTTGAATTCACATTT 559
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 485 CCTGCTGATGTGGAATGGCTTACTGTCCCAATCAAAAGTTCAATTAATCTACTATGT 544
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 560 CCTGCTGATGTGGAATGGCTTACTGTCCCAATCAAAAGTTCAATTAATCTACTATGT 616
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 545 TTGCCAATATGCTCTGCTGTGAATTTGGGCTATATGATATATGCTTATGAACAAG 604
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 617 TTGTCATTAATGTCTGTGGTCAATTAATGTAAGAACTATACCTTACACAGAAAG 676
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 605 AGCACTTGTGCAAGTTGCCAGATTAATCTGTGAGAGTGTGAGTATGACCAATGGTGGCA 664
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 677 AGAAGCTTGTGCAAGTTGCCAGATTAATCTGTGAGAGTGTGAGTATGACCAATGGTGGCA 736
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 665 GTACGAGATCTATTAATTAATCTGTGAAGTTGAAGTCAATTAACCTTAACATCA 724
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 737 ATATGAGATTAATTAATTAATCTGTGAAGTTGAAGTCAATTAACCTTAACATCA 796
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 725 GTTGTGAGGAGACAGTTCAGAGCATCTGCAATTTGTTCAACAGCATTTTAAATACG 784
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 797 ACTTCTTAAGAAAGTTGTCAGAGCTTCAATGCTTCTGTGAACAAATTAATTAATTC 856
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 785 CATACACACGAGAGGAGTATGAGAGAGAGTCAATTAATTAATTAATTAATTAATTC 844
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 857 CAGTCCAA-ATATCAGAGACCATAGAGAAAGAAATACCTCTACTTACTTACTTAT 915
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 845 CTACTTGAATTAACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 904
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 916 GTCC-----CACCAAAATATATGAGTTTGTAGTACTGAATTAATTAATTAATTAAT 965
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 905 TTCAAGTCTTCTGCTGAGATGCTTTTATTTTAAATAATTTTTCATACAAATG 964
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 966 AAGATTTCTGTTCTCTCTCTATTTCTCTATTTTGCATATTAATTAATTAATTAATTAAT 1025
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 965 GTTAAAGAAACAAATCTATTAACCAACACTTGGATTTTATTAATTAATTAATTAATTAAT 1024
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1026 TTTTAAATTAACAAA-----TCAATACCACTTTTGGAACTGGCATATGAATATCTGTA- 1080
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1025 TTAATTTACTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1084
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 -CACATTTATGGAATCAATATCAATATCCACAGATTAAT-ATATTAATTTGTCTATGACTTA 1137
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1085 GTTCACTAAACCTCGATTTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 1144
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1138 AGTCACTAAATCTCTGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1145 GAACA 1149
   |||||
Db 1198 GAACA 1202

```

Search completed: March 14, 2003, 02:07:40
 Job time : 6413 secs